REMARKS

Claims 1-57, all the claims pending in the application, stand rejected on prior art grounds. Claims 1, 20, and 39 are amended herein. Applicants respectfully traverse these rejections based on the following discussion.

I. The Prior Art Rejections

Claims 1-57 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Palmer, et al. (U.S. Patent No. 6,990,628 B1), hereinafter referred to as "Palmer" in view of Woo (U.S. Patent No. 7,039,641 B2). Applicants respectfully traverse these rejections based on the following discussion.

Palmer teaches a method and apparatus for determining when electronic documents stored in a large collection of documents are similar to one another. A plurality of similarity information is derived from the documents. The similarity information may be based on a variety of factors, including hyperlinks in the documents, text similarity, user click-through information, similarity in the titles of the documents or their location identifiers, and patterns of user viewing. The similarity information is fed to a combination function that synthesizes the various measures of similarity information into combined similarity information. Using the combined similarity information, an objective function is iteratively maximized in order to yield a generalized similarity value that expresses the similarity of particular pairs of documents. In an embodiment, the generalized similarity value is used to determine the proper category, among a taxonomy of categories in an index, cache or search system, into which certain documents belong.

Woo teaches a method and system for classifying packets through the use of filters combines heuristic tree search with the use of filter buckets. This provides high performance and reasonable storage requirement, even when applied to large number of filters (from 4K to 1 million). In addition, the method can adapt to the input packet distribution by taking into account the relative filter usage. The capability of employing a large number of filters in a packet classification system is useful in providing value-added services, such as security, quality of service (QoS), load balancing, and traffic accounting.

However, the claimed invention, as provided in amended independent claims 1, 20, and 39 contain features, which are patentably distinguishable from the prior art references of record; namely, Palmer in combination with Woo. For example, claim 1 recites, "[a] system for extracting information comprising: a query input; a database of documents; a plurality of classifiers arranged in a hierarchical cascade of classifier layers, wherein each classifier comprises a set of weighted training data points comprising feature vectors representing any portion of a document, wherein each said feature vector is arranged only as a vector of counts for all features in a data point, and wherein said classifiers are operable to retrieve documents from said database ~~matching~~ based solely on whether said documents are relevant to said query input; and a terminal classifier weighing an output from said cascade according to a rate of success of query terms being matched by each layer of said cascade." Similarly, amended independent claims 20 and 39 recite, "[a]A method of extracting information, said method comprising: inputting a query; searching a database of documents based on said query; retrieving documents from said database ~~matching~~ based solely on whether said documents are relevant to said query using a plurality of classifiers arranged in a hierarchical cascade of classifier layers, wherein

each classifier comprises a set of weighted training data points comprising feature vectors representing any portion of a document, wherein each said feature vector is arranged only as a vector of counts for all features in a data point; and weighing an output from said cascade according to a rate of success of query terms being matched by each layer of said cascade, wherein said weighing is performed using a terminal classifier." Support for the amended language is found in the Applicants' specification, as originally filed (see page 11, lines 12-18 and page 24, line 15 to page 25, line 3 of the Applicants' specification).

The work of Palmer deals with similarities between documents, using specific features such as URLs. Palmer uses a training algorithm to iteratively maximize an objective function of the similarity matrix. Therefore, Palmer requires that the amount of calculation and storage be proportional to the square of the number of documents. Conversely, the Applicants' invention uses feature vectors only (i.e., solely) instead of matrices, which require only resources to be proportional to the number of documents. Moreover, Palmer's classification of documents is performed independent of a specific query. For example, in Palmer the classification occurs, in part, based on the time spent by a user browsing various documents (column 10, lines 15-34). Conversely, the Applicants' invention only performs its classification to determine whether documents are relevant to a specific topic, ignoring their similarities in all other aspects not relevant to the query.

Additionally, Palmer calculates a weight matrix describing similarity information $w(i,j)$ for each pair of documents i and j (column 12, lines 18-19). This is calculated from a number of relationships between these two documents (column 4, line 27 to column 10, line 51). This calculation is performed beforehand, independent of user queries (column 2, lines 64-68, and

column 3, lines 1-3). In contrast, the Applicants' invention only calculates the probability of

relevance of each document to the specific topic of the user query. Accordingly, the Applicants'

invention does not have to calculate the similarity between each pair of documents. In other

words, Palmer's invention does not take into account the user query, and it requires the

availability, calculation, and storage of pairwise information, which is not applicable in the

Applicants' invention. Conversely, the Applicants' invention does not make use of pairwise

information even when it is available.

The work of Woo deals with packet filters for network traffic where a large number of

simple filters are employed to process a large number of packets at a very fast rate. These filters

perform very simple binary decisions. The emphasis in Woo is on getting through as many

filters as possible by by-passing branches of the decision tree. The applicability of Woo's

invention relies on the fact each filter can make deterministic binary decisions. Woo also talks

about weights (column 5, lines 53-67), however the values calculated from these weights are

again used to shape the decision tree (columns 11 and 12); i.e., by-passing branches on the

decision tree.

In contrast, the Applicants' claimed invention propagates probability (feature) vectors

from one classifier (filter) to another (i.e., "each layer of the cascade of classifiers is trained in

succession from a previous layer by the expectation maximization methodology, wherein the

output distribution is used as an input distribution for a succeeding layer."). Moreover, in Woo

results are thresholded to zero or one (i.e., according to claim 5 of Woo "each non-leaf node has

two child nodes, one said child node representing a set of filters having a 0 or * bit at the bit

position corresponding to the non-leaf node, and the other said child node representing a set of

filters having a 1 or * bit at the bit position corresponding to the non-leaf node."). Conversely, in the Applicants' invention, results calculated from each classifier layer are not thresholded to zero or one.

Furthermore, one of the applications for the Applicants' invention is in text analytics, where it is known to those skilled in the art that the thresholding of the results or by-passing of the layers most often degrades the quality of the result. Accordingly, the Applicants' claimed invention is patentably distinct from Woo (in combination with Palmer).

The work of Woo is in a completely different field (i.e., packet filtering in network traffic) from either the Applicants' invention or Palmer's invention (i.e., text information extraction). On the one hand, in Woo the packets arrive in an endless time sequence, and have to be disposed of quickly, independent of each other (column 1, lines 38-47). Neither Palmer's invention nor the Applicants' invention are applicable to this field, since both require an iterative calculation on a fixed large collection of documents. On the other hand, Woo's invention assumes that the filter rules are all fixed beforehand (column 17, lines 44 - 46). Therefore it is not applicable to text information retrieval, where it is important to learn the filter rules (weight matrix in Palmer's invention and probability vectors in the Applicants' invention). Woo's invention bypasses a large number of filters yet still provides an answer as if all the filters have been consulted (column 1, line 66 to column 2, line 2). This is only possible if all the filters make yes-no decisions; conversely in the Applicants' invention the classifiers in text information retrieval keep weights or probabilities which are not zero or one. Since Woo's invention solves problems completely different from the problems solved by either Palmer or the Applicants, the above contrast is made only approximately, assuming that packets correspond to documents and

filters correspond to classifiers, as implicated by the last paragraph of page 3 of the Office Action. In reality, packets are very simple and well-defined entities (see Woo, column 1, lines 26-32), while texts are complex and unstructured (see Palmer, column 4, line 27 to column 10, line 51). Accordingly, Woo in combination with Palmer would result in an inoperable device/method producing conflicting results.

In fact, the USPTO in classifying Woo and Palmer has essentially determined that they are from non-analogous art fields. For example, the USPTO has classified Palmer in U.S. Classes 715/500; 715/501.1; 707/3; and 707/6. Conversely, the USPTO has classified Woo in U.S. Classes 707/100; 370/392; 370/389; and 370/401. Thus, one of ordinary skill in the art would not have been motivated to combine Palmer with Woo especially considering that the USPTO makes no suggestion of such a combination.

The comments on pages 4-5 of the Office Action regarding Applicants' claims 2, 21, 40, 3, 22, 41 suggests that Woo's invention deals with "input distribution" and "output distribution", but there are no such concepts in Woo in the columns/lines cited. Instead, Woo only deals with "input tuples" and "output tuples", which those skilled in the art would readily understand are binary numbers. In contrast, a "distribution" in the context of the Applicants' invention is not a single tuple, but an assignment of probabilities to all possible such tuples (see Applicants' specification, page 13, lines 14-17; page 14, lines 9-21; page 15, lines 13-19). Accordingly, as the MPEP suggests, the Applicants may be their own lexicographers (MPEP §2111.01(III)), and as such the Applicants' claimed language should be read in light of the definitions and descriptions provided in the Applicants' specification.

Furthermore, the comment on pages 6-8 of the Office Action suggests that Woo's

invention discloses a "expectation maximization methodology that maximizes a likelihood of a joint distribution of said training data points and latent variables". The term "expectation maximization algorithm" in the Applicants' invention refers to a specific statistical procedure that is applicable to systems with a specific kind of statistical models (a latent variable model is one of them). There is nothing in Woo (column 9, lines 16-21) that remotely relates to any of this. The following terms are technical terminology related to statistical models: expectation maximization algorithm, likelihood, joint distribution, latent variables. However, these terms do not apply to Woo's setting where a statistical model is absent.

Moreover, the comment on page 4 of the Office Action about "the same field of endeavor of database management system, such as, searching, classifying data, weights and frequencies" suggests a possible misunderstanding of the terminology. Database systems store data and allow the user to change and search/retrieve them later. However, a packet filtering system, such as in Woo, makes decisions on each packet and promptly forgets about them. Searching is only possible if data is stored. The "frequency" in Woo is the physical frequency, which refers to how many packets arrive in a period of time, while the "frequency" as used by the Applicants' is the statistical frequency, such as how many times a word appears in a document.

In Woo, frequency is an operational requirement, while in the Applicants' invention, frequency is one of the features used in the calculation. In a non-technical sense, all three systems can be loosely said to "classify data". However, the Federal Circuit cautions not to read claims in a vacuum, but rather in the context of the specification. In re Marosi, 710 F.2d 799, 802 218 USPQ 289, 292 (Fed. Cir. 1983) (quoting In re Okuzawa, 537 F.2d 545, 548, 190 USPQ 464, 466 (CCPA 1976)). Thus, the Applicants' claims must be in light of the language in the

specification.

Generally, Woo makes a quick decision on packets as they arrive, with predefined rules; Palmer classifies all documents into categories and calculates similarities to be used later; while the Applicants' invention classifies documents relative to a particular user query. The result of the classification in both Palmer and the Applicants' invention depends on what other documents are in the collection, while the concept of collection is absent from Woo. Woo uses weights not to make decisions, but rather to choose decision makers (filters) so as to save time; the final decision should not depend on the weights if the invention in Woo is functionally correct and operable. Accordingly, the Applicants' invention provides a manner of combining the probability vector at each layer to form an overall probability vector that can be used to make final classification. Neither Palmer nor Woo teaches this.

Insofar as references may be combined to teach a particular invention, and the proposed combination of Palmer with Woo, case law establishes that, before any prior-art references may be validly combined for use in a prior-art 35 U.S.C. § 103(a) rejection, the individual references themselves or corresponding prior art must suggest that they be combined.

For example, in In re Sernaker, 217 USPQ 1, 6 (C.A.F.C. 1983), the court stated: "[P]rior art references in combination do not make an invention obvious unless something in the prior art references would suggest the advantage to be derived from combining their teachings." Furthermore, the court in Uniroyal, Inc. v. Rudkin-Wiley Corp., 5 USPQ 2d 1434 (C.A.F.C. 1988), stated, "[w]here prior-art references require selective combination by the court to render obvious a subsequent invention, there must be some reason for the combination other than the hindsight gleaned from the invention itself. . . . Something in the prior art must suggest the

desirability and thus the obviousness of making the combination."

In the present application, the reason given to support the proposed combination is improper, and is not sufficient to selectively and gratuitously substitute parts of one reference for a part of another reference in order to try to meet, but failing nonetheless, the Applicant's novel claimed invention. Furthermore, the claimed invention, as amended, meets the above-cited tests for obviousness by including embodiments such as "retrieving documents from said database based solely on whether said documents are relevant to said query" and "wherein each said feature vector is arranged only as a vector of counts for all features in a data point." As such, all of the claims of this application are, therefore, clearly in condition for allowance, and it is respectfully requested that the Examiner pass these claims to allowance and issue.

As declared by the Federal Circuit:

In proceedings before the U.S. Patent and Trademark Office, the Examiner bears the burden of establishing a prima facie case of obviousness based upon the prior art. The Examiner can satisfy this burden only by showing some objective teaching in the prior art or that knowledge generally available to one of ordinary skill in the art would lead that individual to combine the relevant teachings of the references. In re Fritch, 23 USPQ 2d 1780, 1783 (Fed. Cir. 1992) citing In re Fine, 5 USPQ 2d 1596, 1598 (Fed. Cir. 1988).

Here, the Examiner has not met the burden of establishing a prima facie case of obviousness. It is clear that, not only does Palmer fail to disclose all of the elements of the claims of the present invention, particularly, "retrieving documents from said database based solely on whether said documents are relevant to said query" and "wherein each said feature vector is arranged only as a vector of counts for all features in a data point," as discussed above, but also, if combined with Woo, fails to disclose these elements as well. The unique elements of the claimed invention are clearly an advance over the prior art.

10/723,112   22

The Federal Circuit also went on to state:

The mere fact that the prior art may be modified in the manner suggested by the Examiner does not make the modification obvious unless the prior art suggested the desirability of the modification. . . . Here the Examiner relied upon hindsight to arrive at the determination of obviousness. It is impermissible to use the claimed invention as an instruction manual or "template" to piece together the teachings of the prior art so that the claimed invention is rendered obvious. This court has previously stated that one cannot use hindsight reconstruction to pick and choose among isolated disclosures in the prior art to deprecate the claimed invention. Fritch at 1784-85, citing In re Gordon, 221 USPQ 1125, 1127 (Fed. Cir. 1984).

Here, there is no suggestion that Palmer, alone or in combination with Woo teaches a method and structure containing all of the limitations of the claimed invention. Consequently, there is absent the "suggestion" or "objective teaching" that would have to be made before there could be established the legally requisite "prima facie case of obviousness."

In view of the foregoing, the Applicants respectfully submit that the cited prior art references of record do not teach or suggest the features defined by amended independent claims 1, 20, and 39 and as such, claims 1, 20, and 39 are patentable over Palmer in combination with Woo. Further, dependent claims 2-19, 21-38, and 40-57 are similarly patentable over Palmer in combination with Woo, not only by virtue of their dependency from patentable independent claims, respectively, but also by virtue of the additional features of the invention they define. Moreover, the Applicants note that all claims are properly supported in the specification and accompanying drawings. In view of the foregoing, the Examiner is respectfully requested to reconsider and withdraw the rejections.

II.   **Formal Matters and Conclusion**

In view of the foregoing, Applicants submit that claims 1-57, all the claims presently

pending in the application, are patentably distinct from the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary. Please charge any deficiencies and credit any overpayments to Attorney's Deposit Account Number 09-0441.

Respectfully submitted,

Dated: <u>August 7, 2006</u>

Mohammad S. Rahman
Registration No. 43,029

Gibb I.P. Law Firm, LLC
2568-A Riva Road, Suite 304
Annapolis, MD 21401
Voice: (301) 261-8625
Fax: (301) 261-8825
Customer Number: 29154